

Blockchains and Distributed Markets for Library Metadata

Timothy A. Thompson
Yale University Library
timothy.thompson@yale.edu

Anyone familiar with library or information science probably knows something about the discourse of control. Authority control, controlled vocabularies, intellectual and physical control of collections—the idea of control is part of a professional ethos that tends to privilege centralization and uniformity in the organization of information. Centralization on the level of organization and representation also maps to centralization on the level of maintenance and exchange: libraries rely on gatekeepers and centralized services to facilitate initiatives in cooperative cataloging and to manage resources such as the Library of Congress Name Authority File and knowledge organization systems such as the Library of Congress Subject Headings. Ultimately, questions of control are questions of trust: centralized systems and workflows allow libraries to define boundaries within which community norms and standards can be enforced. At the same time, they are exclusionary and limit participation to a set of authorized contributors. Blockchain technology has the potential to reconfigure relations of information exchange among libraries and to shift the locus of trust from centralized services to distributed systems built on the premise of peer-to-peer interaction.

As an illustration, consider the example of the NACO–México project [1], which began in 2003 as a cooperative effort among academic libraries in Mexico to contribute records to the Name Authority Cooperative Program (NACO), managed by the Program for Cooperative Cataloging and the Library of Congress. Currently, any institution that wishes to contribute to the Name Authority File must do so through one of two authorized gateways, both of which impose membership fees. In 2011, participants in NACO–México were compelled to move from the primary NACO gateway to its alternative because they could no longer afford the cost of membership in the former. More recently, the project has dealt with discrepancies in its contribution statistics as recorded by the Library of Congress, which has reported contribution totals that are lower than the project's internal numbers. If NACO itself were to run on a public blockchain rather than as a centralized service, membership fees might be replaced by marginal transaction fees, making participation affordable for international participants who were previously unable to devote resources to funding and coordinating the process of contributing data. Furthermore, discrepancies in contribution statistics could be eliminated through reference to the blockchain as a permanent, timestamped record of transactions. Individual libraries might purchase data on demand from peer contributors, decreasing local workloads and creating new opportunities for international collaboration.

For several years, academic libraries have been discussing and exploring efforts to migrate from legacy metadata formats to linked open data. The transition has been a difficult one. Although work developing semantic ontologies such as BIBFRAME has allowed the library community to examine and reconsider some of its fundamental data models, the implementation of linked open data in academic libraries has been impeded by the absence of an underlying computational architecture that can support new models for sharing and production. A major selling point of linked open data is its support for internationalization and integration with the wider web of data. However, without an open, distributed market for the exchange of data, centralized bottlenecks will continue to undermine attempts at systemic change. By providing an international marketplace for peer-to-peer exchange, blockchains could facilitate the free flow of library metadata, potentially creating new revenue streams for individual libraries and replacing some of the costly subscription

services that currently predominate. In short, the adoption of public blockchains could give libraries a new kind of control over their metadata, the control of ownership and autonomy.

[1] J. M. Martínez Saldaña, "Informe sobre las actividades del proyecto NACO-México," presented at Sites/Cites, Texts, and Voices in Critical Librarianship: Decolonizing Libraries and Archives, Seminar on the Acquisition of Latin American Library Materials (SALALM) 63, Mexico City, Mexico, July 2, 2018.